Reviewer's report

Title: Multidetector CT Angiography for assessment of in-stent restenosis: Meta-analysis of Diagnostic Performance

Version: 1 Date: 6 January 2008

Reviewer: Hans Reitsma

Reviewer's report:

Major compulsory revisions

The authors performed a systematic review, including a meta-analysis, of Mutidetector CT angiography to detect in-stent restenosis. Here is a list of my main comments paying special attention to the statistical analysis.

Framing the question

For a systematic review to be relevant and informative it is critical to start with well-formulated questions. Although the authors provide information about several key elements like the test under evaluation (index test), reference standard and patient population, they do not discuss the intended use or future role of the test under evaluation. Being specific about the future role is important as it will provide the necessary background information: to judge whether a comparative review would be more informative; to choose the most informative accuracy measure when reporting and interpreting the results; and it provides the framework against which to discuss whether accuracy is high enough in relation to the intended role of the index test.

Key elements to address are:

+ Are there other non-invasive techniques that can measure in-stent restenosis and is the real interest whether non-invasive technique A is better than technique B (comparative review)?
+ Is the intended role of the index test one of replacement of a other (prior) test or should it replace the reference standard (replacement)?
+ Is the intended use of the index test to select patients that will go on to receive the (invasive) reference standard or not (triaj tests) See also paper by Bossuyt et al. BMJ 2006:332:1089-1092.

For instance, if the intended role is that of a triage test, the number of false negative index test results is of primary concern, making sensitivity the key outcome to consider. If replacement is the issue to consider, the required height of sensitivity and specificity should be discussed in relation to the consequences of false positive and false negative test result and the frequency and severity of side-effects associated with the invasive reference standard procedure.

Search strategy
The search strategy is not completely described in this paper. It appears that only MeSH have been used, but previous studies have shown that adding free text words can identify additional relevant articles.

Quality assessment

The authors have performed a formal quality assessment of included studies using the QUADAS instrument. Unfortunately, they then proceed to calculate a summary quality score. Summary quality scores have been extensively studied and criticized both in systematic reviews of intervention and diagnostic studies. The main reason is that it impossible to determine the weight that should be given to each quality item because the relevance of an item may differ depending on the type of index test and the type of design (see also Juni et al. JAMA 1999;282:1054-1060 and Whiting et al. BMC Med Res Methodol 2005;5:19). The results of the quality assessment (individual items) are not given in the main paper.

Meta-analysis and presenting the data

I have several objections in relation to the statistical analyses.

+ Choice of accuracy measure

The authors have meta-analysed all available accuracy measures in their review which in my view is not helpful to the reader. They should restrict the reporting to those statistics that are most informative in relation to the intended role of the test. Furthermore, the methods that have been used for some of these are not statistically sound (see statistical pooling).

+ Assessment of publication bias

Publication bias was investigated according to the method of Sterne and Egger. These methods can be seriously misleading in diagnostic systematic reviews (see Deeks et al. J Clin Epidemiol 2005;58:882-893). The reason is that the traditional funnel plot (plotting log diagnostic odds ratio against SE) can show asymmetry in situations where there is no publication bias. Furthermore, although the authors find that there is substantial departure from the ideal funnel shaped distribution in their review, the authors do not discuss the implications of their finding.

+ Statistical pooling

Estimates of sensitivity and specificity were pooled separately, but this ignores the fact that sensitivity and specificity might be negatively correlated. Two more advanced methods exist that properly take into account any possible correlation, in addition to within-study variation (precision) and between-study variation (random effect approach): these are the bivariate meta regression approach (Reitsma et al. J Clin Epidemiol 2005;58:982-990) and the hierarchical summary ROC approach (Rutter et al. Stat Med 2001;20:2865-2884). These methods can also be extended to incorporate covariates to examine how accuracy changes according to specific factors. The standard method of Moses-Littenberg can be used as a descriptive tool to obtain a summary ROC curve, but recent work has
shown that the method can perform poorly when calculating p-values and confidence intervals. Therefore, it should not be used in these situations.

The authors also pool positive and negative likelihood ratios, although this can produce misleading results (Why not to pool likelihood ratios. Zwinderman AH, Bossuyt PM. Stat Med 2007;). It would have been better to calculate them from the pooled estimates of sensitivity and specificity.

Interpreting the results / Discussion
It is often useful to make the results of the diagnostic meta-analyses more accessible by using absolute numbers. One way to do this is to use the summary estimates of sensitivity and specificity and a reasonable estimate for the prevalence (or a range of prevalence reflecting different settings) and calculate the number of true positives, false positives, true negatives, false negatives (or related predictive measures) based on a round number of patients (for instance 1,000).

Minor essential revisions
+ Confidence intervals are missing in the Abstract
+ For several factors it is mentioned that they had an impact on accuracy but the direction and/or magnitude of the effect is not given (both Abstract and Results)

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests